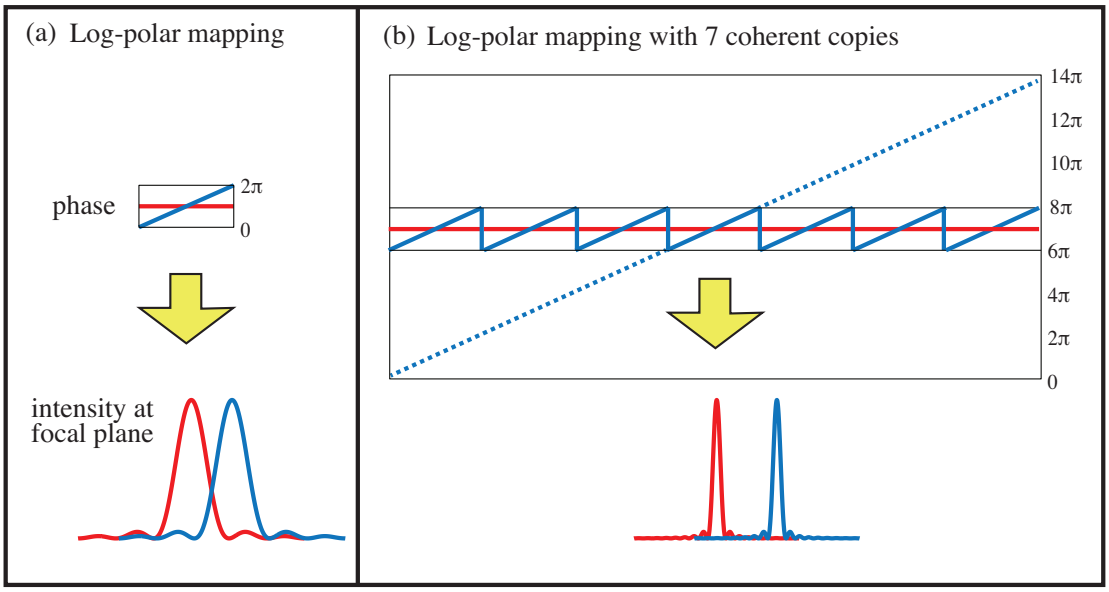
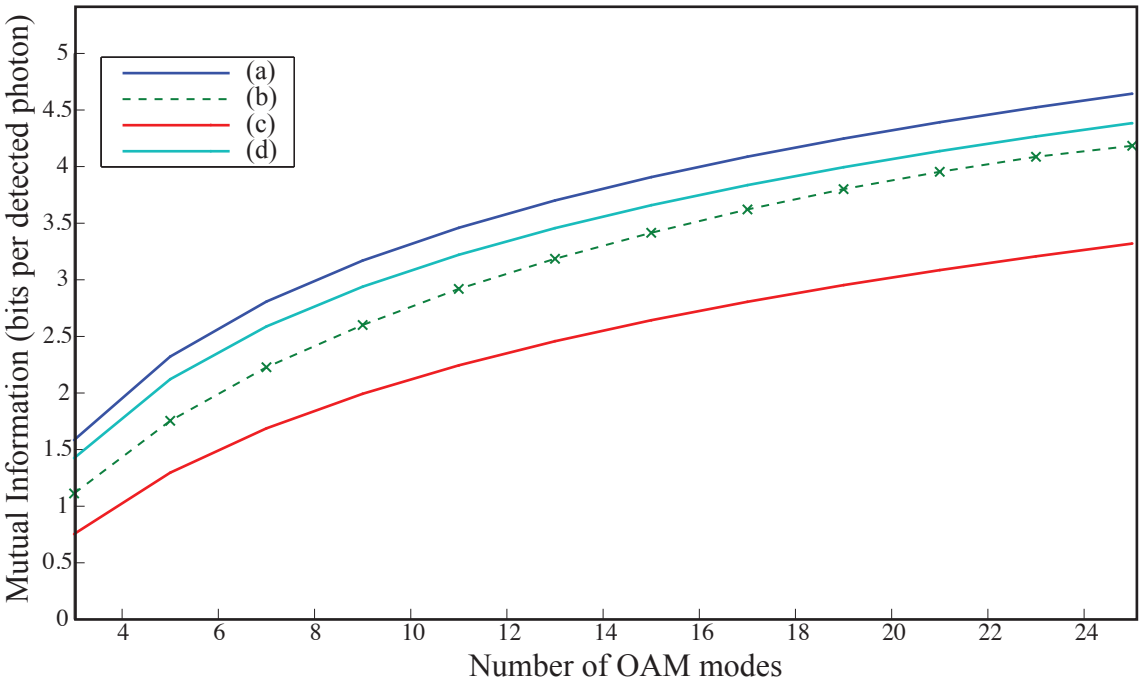
Supplementary Figure S1



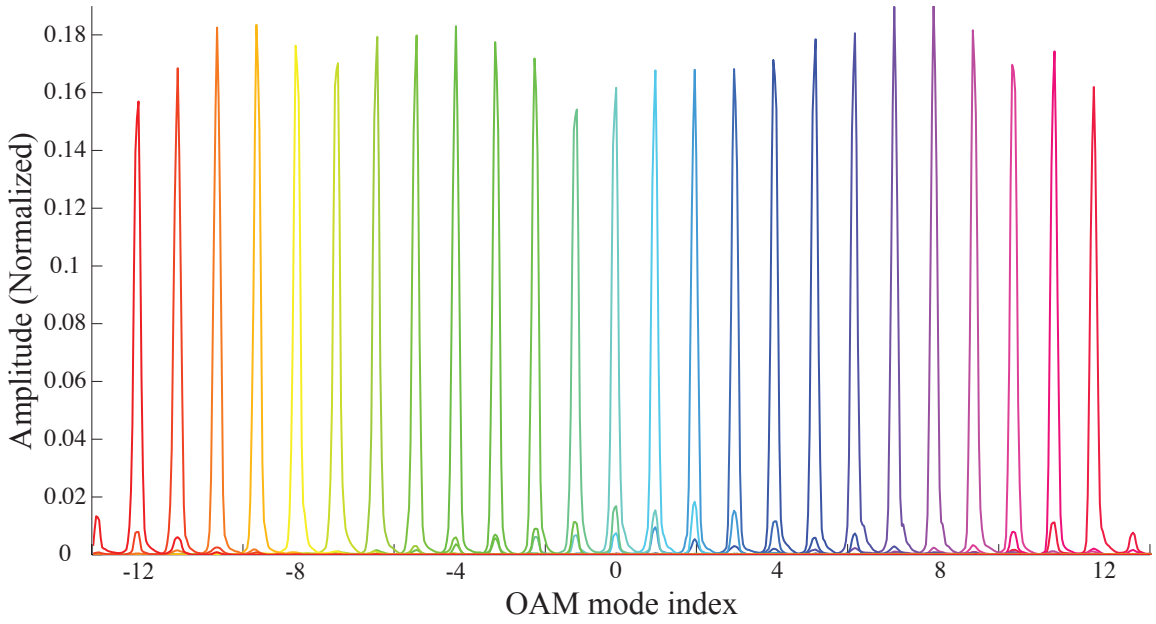
Supplementary Figure S1: Mode overlap between transformed OAM modes. The phase structure of two neighboring OAM modes and the resulting intensity patterns in the Fourier domain for **a)** perfect log-polar mapping, and **b)** perfect log-polar mapping combined with beam copying.

Supplementary Figure S2



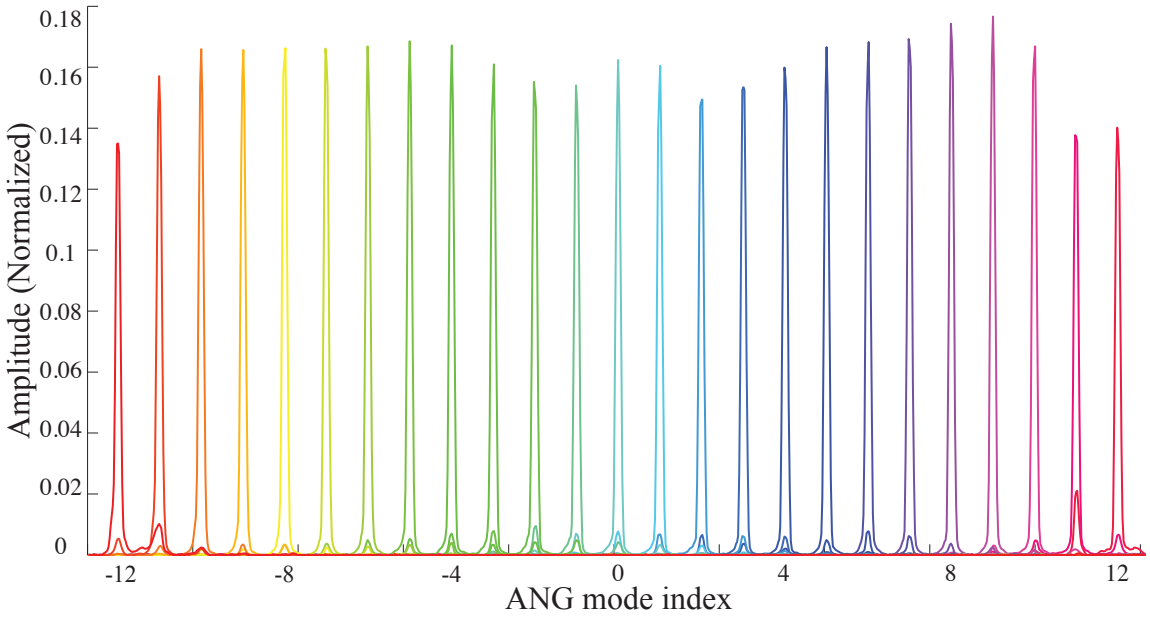
Supplementary Figure S2: Mutual information as a function of the number of the modes. Mutual information as calculated for **a)** ideally separated modes, **b)** data from this experiment, **c)** perfect log-polar mapping, and **d)** perfect log-polar mapping combined with beam copying.

Supplementary Figure S3



Supplementary Figure S3: Intensity profile of the transformed OAM modes. The incident power in each OAM mode (summed along the vertical axis) as a function of the horizontal coordinate on the CCD camera. The center of each mode is labelled with the corresponding OAM mode index. The sum of the total power in each mode is normalized to unity.

Supplementary Figure S4



Supplementary Figure S4: Intensity profile of the transformed ANG modes. The total incident power in each ANG mode as a function of the horizontal coordinate on the CCD camera. The center of each mode is labelled with the corresponding ANG mode index. Sum of the total power in each mode is normalized to unity.

Supplementary Table S1

Diffraction order m	-3	-2	-1	0	1	2	3
$\gamma_7(m)$	1.24	1.45	1.28	1	1.28	1.45	1.24
$\alpha_7(m)$ (rad)	7.03	1.89	-0.99	0	-0.99	1.89	7.03

Diffraction order m	-4	-3	-2	-1	0	1	2	3	4
$\gamma_9(m)$	1.03	0.943	0.963	0.971	1	0.971	0.963	0.943	1.03
$\alpha_9(m)$ (rad)	1.41	3.03	5.57	0.72	0	0.72	5.57	3.03	1.41

Supplementary Table S1: Parameters for fan-out elements. (top) Grating for creating seven copies. **(bottom)** Grating for creating nine copies.

Supplementary Note 1. Optimal Designs for Refractive Beam Copying

Phase-only elements for beam copying can be designed using phase-retrieval and other iterative algorithms [25]. Alternatively, an analytical optimization process can be used to design such elements [26]. The results from these two methods agree closely. These elements have the general form

$$\Psi_{2N+1}(x) = \tan^{-1}\left(\frac{\sum_{m=-N}^{N} \gamma_m \sin[(2\pi s/\lambda)mx + \alpha_m]}{\sum_{m=-N}^{N} \gamma_m \cos[(2\pi s/\lambda)mx + \alpha_m]}\right). \tag{S1}$$

The values of γ and α for fan-out elements creating 7 and 9 copies are listed in Supplementary Table S1.

Supplementary Note 2. Separation Efficiency

We define the separation efficiency as the probability of detecting an OAM mode correctly. The mode sorter maps different OAM modes to a series of spots which their centroids separated. Assuming a perfect log-polar mapping, each OAM mode is mapped to a plane-wave which has a finite size [12].

$$E(v) = e^{i\ell v/a}\mathrm{rect}(\frac{v}{2\pi a}). \tag{S2}$$

The separability of the form of these modes allows for a one-dimensional analysis. This unwrapped mode is focused to a spot. Using basic Fourier analysis, the spot pattern is found to be

$$E(x) \propto \frac{\sin(\pi[\frac{ax}{\lambda f} - \ell])}{\frac{\pi x}{\lambda f}}, \tag{S3}$$

where f is the focal length of the lens. Two transformed neighboring OAM modes are orthogonal in terms of their amplitude due the unitary nature of the involved transformation. The amplitude profile of these modes , however, is the familiar sinc form, which has a strong central components with side lobes around it (See Supplementary Fig. S1). As a result, there is an overlap between two neighboring OAM modes.

The probability of detecting each mode is equal to the portion of the intensity which falls into the corresponding spatial bin for an OAM mode. To treat all the OAM modes equally, the spatial bins have to be confined by the borders between the neighboring modes. Using the above analysis, the separation efficiency can be calculated as

$$\text{S. E.} = \frac{\int_{-1/2}^{1/2} \left|\frac{\sin(\pi x)}{\pi x}\right|^2 dx}{\int_{-\infty}^{\infty} \left|\frac{\sin(\pi x)}{\pi x}\right|^2 dx} = 0.774 = 77.4\%. \tag{S4}$$

When a fan-out element is added, the unwrapped OAM mode is copied to N different copies. For an odd number of copies

$$E(v) = \frac{1}{\sqrt{N}} e^{i\ell v/a} \sum_{i=-(N-1)/2}^{(N-1)/2} \mathrm{rect}(\frac{v - ia}{a}) = \frac{1}{\sqrt{N}} e^{i\ell v/a}\mathrm{rect}(\frac{v}{Na}). \tag{S5}$$

The factor of $1/\sqrt{N}$ has to be added for conservation of energy. This extended unwrapped mode is focused to a spot which is N times narrower

$$E(x) \propto \frac{\sqrt{N}\sin(\pi[\frac{Nxa}{\lambda f} - \ell])}{\frac{\pi x}{\lambda f}}. \tag{S6}$$

In this case, the sinc functions are N times narrower as compared to the previous case but the spacing between them has not changed (See Supplementary Fig. S1). We can use an analysis similar to above to calculate the separation efficiency for the case where the log-polar mapping is enhanced with beam copying. For the case of seven copies, the separation efficiency is equal to

$$\text{S. E.} = \frac{\int_{-1/2}^{1/2} \left| \frac{\sin(7\pi x)}{7\pi x} \right|^2 dx}{\int_{-\infty}^{\infty} \left| \frac{\sin(7\pi x)}{7\pi x} \right|^2 dx} = 0.974 = 97.4\%. \tag{S7}$$

Supplementary Note 3. Effects of Cross-talk and Loss on Mutual Information

A common measure to quantify the degree of separation of modes in a communication system is the mutual information between the sent and received modes. This quantity is defined as

$$I(Y,X) = \sum_{i,j} P(y_j, x_i) \log_2 \left[\frac{P(y_j, x_i)}{P(x_i)P(y_j)} \right]. \tag{S8}$$

Here, x_i is the event of sending a photon with the OAM value of ℓ_i and y_j is the event of detecting a photon with the OAM value of ℓ_j. Assuming a uniform probability for sending N modes, P_x equals $1/N$ for all x. The joint probability of detection $P(y_j, x_i)$ can be calculated from the conditional probability $P(y_j|x_i)$ by using the Bayes' rule

$$P(y_j, x_i) = P(y_j|x_i)P(x_i) = \frac{1}{N} P(y_j|x_i). \tag{S9}$$

We have spatially binned the detector plane in order to detect each mode. In this case, the conditional probability $P(y_j|x_i)$ is proportional to incident power from mode i to the bin corresponding to the mode j (For patterns of the detected modes, see Supplementary Fig. S3 and Supplementary Fig. S4). The proportionality constant can be simply found by imposing the probability normalization condition $\sum_j P(y_j|x_i) = P(x_i) = \frac{1}{N}$.

The normalization condition described above excludes the loss in the channel (we denote the conditional probabilities measured from this method as $P(y_j|x_i) = S_{ji}$). As a result, the mutual information calculated from this method only quantifies the degree of cross-talk and it is usually measured in units of bits per detected photon. Since our method relies on phase-only optical elements, in principle it can achieve a power transmission of unity. It should be noted that although the fan-out elements have a smaller-than-unity efficiency (99.3 for 9 copies and 96.7 for 7 copies) when considering only the desired copies, they are phase-only components and are strictly lossless. In our experiment and the theory work published previously [24] we collect all the light diffracted from the fan-out element and as a result there is no inherent loss in our method.

Using the values of S_{ji} matrix, the mutual information between the transmitted and the detected modes can be calculated as

$$I(\text{bits per detected photon}) = \frac{1}{N} \sum_{i,j=1}^{N} S_{ji} \log_2 \left[\frac{N S_{ji}}{\sum_i S_{ji}} \right]. \tag{S10}$$

We have used the conditional probability matrix S_{ji} to calculate the mutual information for OAM basis sets with different dimensions (Supplementary Fig. S2). For comparison, we have also plotted the theoretical

upper bound for the mutual information of a system employing N OAM modes, which equals $\log_2(N)$ (Supplementary Fig. S2). Supplementary Fig. S2 also includes values of mutual information calculated from theory for a perfect log-polar coordinate mapping [12] (c) and perfect log-polar mapping combined by beam copying (d) [24] (all values are in bits per detected photons).

To achieve a complete characterization of the information capacity of a channel, one must take into account the loss as well as the cross-talk in the calculation of the mutual information between the transmitted and the detected modes. Due to the presence of loss in a realistic communication channel, sometimes the transmitted photons lead to no detection event in the receiver's side. In this situation, the failure to detect any mode should be treated as an alternative possible event for the receiver [29]. However, it is easy to check that if the value of the loss is equal for all the modes, the no-detection event contains no information about the transmitted mode and hence it can be discarded from the calculation of the mutual information. In this case, the only effect of loss is in reducing the probability of successfully detecting each mode.

If the transmitted photon is lost in the channel by a probability of P_L, a successful transmission will have a probability of P_T, where $P_T = 1 - P_L$. In this case, the joint probability of detection $P(y_j, x_i)$ for each mode is calculated as

$$P(y_j, x_i) = P(y_j|x_i)P(x_i) = \frac{1}{N} P_T S_{ji}, \tag{S11}$$

where S_{ji} are measured as before. Substituting the joint probability values in Eq. (S8) we have

$$I(\text{bits per launched photon}) = \frac{1}{N} \sum_{i,j=1}^{N} P_T S_{ji} \log_2 \left[\frac{N \cancel{P_T} S_{ji}}{\sum_i \cancel{P_T} S_{ji}} \right]. \tag{S12}$$

This can be further simplified as

$$I(\text{bits per launched photon}) = P_T \times I(\text{bits per detected photon}). \tag{S13}$$

This relation has a simple interpretation. It states that the mutual information between the transmitted and the detected modes is scaled down by the the overall power transmission efficiency of the optical communication link.

We have measured the overall power efficiency of our system to be about 15 percent. Using this number, our system transfers 4.18 bits per detected photon and $4.18 \times 0.15 = 0.63$ bits per launched photon for OAM modes. Similarly, the ANG modes are transferred and separated with a mutual information of 4.16 bits per detected photon and $4.16 \times 0.15 = 0.62$ bits per launched photon. However, it should be emphasized once more that the loss in our system is a consequence of the specific experimental realization we have utilized, and

is not inherent to our sorting method. Straightforward technological improvements could lead to a significant increase in the number of bits per launched photon. Manufacturing of the fan-out and its phase corrector in the form of refractive elements can achieve power transmission efficiencies close to a hundred percent. In fact, fan-out elements with efficiencies of more than 95% are commercially available. Alternatively, the required elements can be realized using new generation of SLMs which can achieve diffraction efficiencies of more than 90%.

Supplementary Note 4. Mutually Unbiased Bases and Angular Modes

To guarantee security in the BB84 protocol, at least two mutually unbiased bases (MUBs) of polarization are needed [4]. Similarly, a basis unbiased with respect to the basis of orbital angular momentum (OAM) is needed in an OAM-based quantum key distribution (QKD) system [10]. Such a basis can be constructed using the following formula

$$|\theta_j\rangle = \frac{1}{\sqrt{2L+1}} \sum_{\ell=-L}^{L} |\ell\rangle \, e^{-i2\pi j\ell/(2L+1)}. \tag{S14}$$

These modes are referred to as the angular (ANG) modes due to the angular confinement of their intensity pattern. Some of the ANG modes made from superposition of $\ell = -12$ to $\ell = -12$ can be seen in Fig. 4.

We have used a phase-only spatial light modulator (SLM) to realize holograms for generating OAM and ANG modes. In order to generate a spatial mode described by $a(x, y)e^{i\phi(x,y)}$, the following phase pattern is impressed on the SLM

$$\Psi(x, y) = \big(1 - \mathrm{sinc}^{-1}[a(x, y)]\big)\, \phi(x, y). \tag{S15}$$

A blaze function is usually added to the phase of the target beam to achieve better separation of the desired and undesired parts in the Fourier domain [27]. Fig. 4 shows some of the holograms used for generating OAM and ANG modes.

Supplementary References

[29] Cover, T. M. & Thomas, J. A. *Elements of Information Theory.* 1st Edition. New York: Wiley-Interscience, 1991.